

---

**Cultural conflict and merger failure:  
An experimental approach**

**Roberto A. Weber and Colin F. Camerer**

**USC Center for Law, Economics & Organization  
Research Paper No. C01-10**



**CENTER FOR LAW, ECONOMICS  
AND ORGANIZATION  
RESEARCH PAPER SERIES**

**Sponsored by the John M. Olin Foundation**

University of Southern California Law School  
Los Angeles, CA 90089-0071

*This paper can be downloaded without charge from the Social Science Research Network  
electronic library at [http://papers.ssrn.com/abstract\\_id=276288](http://papers.ssrn.com/abstract_id=276288)*

---

# Cultural conflict and merger failure: An experimental approach

Roberto A. Weber<sup>1</sup>

Department of Social and Decision Sciences  
Carnegie-Mellon University  
Pittsburgh PA 15213  
[rweber@andrew.cmu.edu](mailto:rweber@andrew.cmu.edu)

Colin F. Camerer

Division of Humanities and Social Sciences 228-77  
California Institute of Technology  
Pasadena CA 91125 USA  
[camerer@hss.caltech.edu](mailto:camerer@hss.caltech.edu)

## Abstract

We use laboratory experiments to explore merger failure due to conflicting organizational cultures. We introduce a laboratory paradigm for studying organizational culture that captures several key elements of the phenomenon. In our experiments, we allow subjects in “firms” to develop a culture and then merge two firms. As expected, performance decreases following the merging of two laboratory firms. In addition, subjects over-estimate the performance of the merged firm and attribute the decrease in performance to other subjects rather than to situational difficulties created by conflicting culture.

---

<sup>1</sup> This work has benefited greatly from the authors’ discussions with Linda Argote, Linda Babcock, Paul Fischbeck, Chip Heath, Herb Simon, and George Loewenstein. Thanks to participants at the 2000 Behavioral Decision Research in Management meetings in Tucson, the 2000 Economic Science Association meetings in New York, and seminar participants at Carnegie Mellon University and Harvard Business School for helpful comments and suggestions. This research was supported by NSF 97-30364.

## Introduction

A majority of corporate mergers fail. Failure occurs, on average, in every sense: acquiring firm stock prices tend to fall slightly when mergers are announced; many acquired companies are later sold off; and profitability of the acquired firm is lower after the merger (relative to comparable non-merged firms).<sup>2</sup> Participants report a lot of conflict during the merger, resulting in high turnover (Buono, Bowditch and Lewis, 1985; Walsh, 1988).<sup>3</sup> Participants express disappointment in the mergers' results, and surprise at how disappointed they are. Perhaps the only players who are uniformly happy are investment bankers, who get paid for putting deals together and get paid again for undoing them. Curiously, widespread merger failure is at odds with the public and media perception that mergers are grand things that are almost sure to create enormous business synergies that are good for employees, stockholders, and consumers.

This paper introduces a simple experimental paradigm to explore possible causes of merger failure. The guiding hypothesis is that an important component of failure is conflict between the merging firms' cultural conventions for taking action, and an underestimation by merger partners of how severe, important, and persistent conflicts are. Cultural conventions emerge to make individual firms more efficient by creating a shared

---

<sup>2</sup> The most conclusive evidence of lower post-merger profitability comes from studies by Ravenscraft and Scherer (1987,1989). They use FTC line-of-business data to compare companies' lines of business after they were acquired with a proxy for what their performance would have been without the merger (using comparable control businesses which were not involved in acquisitions). Operating income as a percentage of assets is lower by 0.03 for the merged target businesses. This is a substantial (and statistically significant) drop, since their pre-takeover operating income/asset ratio averaged 0.115. Also, McGuckin, Nguyen and Reznick (1995) provide support for the hypothesis that mergers and acquisitions fail on average, even though their overall interpretation is the opposite (but not entirely supported by their analysis). Specifically, they find that acquisitions decrease productivity and employment at the firm level (even though acquiring firms were typically highly productive before the acquisition) and this is similarly supported in their initial plant-level analysis. They manage to overturn the productivity result at the plant level only for a subset of plants (those belonging to larger firms).

understanding that aids communication and action.<sup>4</sup> However, when two joined firms differ in their conventions, this can create a source of conflict and misunderstanding that prevents the merged firm from realizing economic efficiency.

In addition, we hypothesize that the extent of these conflicts are unexpected because observers focus on tangible aspects of firms' practices (such as technology and capital) and ignore aspects that are more difficult to measure, such as culture. This leads to overestimation of the value of a merged firm at the time of the merger.

Our emphasis on cultural conflict is not meant to suggest, of course, that other potential causes of merger failure are not important. Certainly, agency problems, optimism, and hubris may lead top managers to undertake mergers that are bad for shareholders. Also, holding cultural incompatibility aside, conflicts of interest between employees in two merged firms may also harm the merger. For instance, employees in each of the two firms may have reasons to prefer maintaining the "old way of doing things" – possibly because of learning costs, inertia, etc. – and may therefore intentionally resist adopting the other firms' practices.<sup>5</sup> While we recognize these other potential sources of merger failure, our focus is on one specific cause: differences in culture may simply make it difficult for members of the merged organization to see things in the same way.

In our experiments, culture is a specialized homemade language a pair develops to solve a task. In the task, two subjects with the same set of pictures have to learn to jointly

---

<sup>3</sup> Walsh and Ellwood (1991) find that the high rate of turnover among management at acquired firms is not related to poor prior performance, indicating that the turnover is not due to the pruning of under-performing management at the acquired firm.

<sup>4</sup> See Schall, 1983.

<sup>5</sup> In a game theoretic sense, this means that employees in the merged firms are playing something more like battle-of-the-sexes or chicken (where players have incentives that are not perfectly aligned), rather than just a pure Schelling-type matching game (where interests are perfectly aligned).

identify a subset of the entire set of pictures. To do this, they must create a shared way to describe the pictures, as quickly as possible, so that a “manager” subject can guide the “employee” to pick the pre-specified subset. Two pairs of subjects, or “firms,” develop cultures separately. Then the two pairs are “merged.” One manager must then describe an object set to two employees at once. Since the manager only participated with one of the two employees before, we anticipated that the conflict in homemade languages would make it difficult for the manager to quickly get the new employee to pick the right pictures, and would **also** slow down speed of the old employee.

Therefore, we expect and find that performance post-merger is significantly affected by the difference in languages. Also, given our hypothesis that participants in a merger typically focus on the gains to be obtained (because of synergies, technology, etc.) and neglect potential cultural integration difficulties, we elicit expectations from subjects about the performance of the merged firm. Consistent with our expectation, subjects over-estimate the value of the firm. Finally, after experiencing integration difficulties, subjects do not attribute the source of these difficulties entirely to difference in languages/cultures, but instead also place some of the blame on people from the other firm.

## **Organizational culture**

Organizational culture has received considerable attention from organizational researchers, and substantially less attention from economists. While agreement on a precise definition of the concept has proven difficult, there are a few important elements shared by most definitions. Culture is usually thought of as a general shared social

understanding, resulting in commonly held assumptions and views of the world among organizational members (Wilkins and Ouchi, 1983; Schall, 1983; Schein, 1983; Rousseau, 1990). Culture is developed in an organization through joint experience, usually over long periods of time. It is useful to an organization because it allows its members to coordinate activity tacitly without having to reach agreement explicitly in every instance. Language – in the form of codes, symbols, anecdotes, and rules about appropriate statements – plays an important role in organizational culture, constituting a large part of the shared understanding held by organizational members (Schall, 1983; Schein, 1983).

However, despite agreement that culture is important, it is difficult to measure and study precisely (Schein, 1996; Marcoulides and Heck, 1993; Rousseau, 1990).

Researchers have relied on a few different approaches to study culture in organizations (see Schein; 1990, Rousseau, 1990). Much of this research is ethnographic observation of interactions in small numbers of organizations (e.g., Schein, 1983 and 1990; Barley, 1983). While informative and helpful for inspiring theory, the small samples involved in this type of analysis usually makes it difficult to draw firm conclusions based on this research.

Another approach consists of questionnaires administered to large numbers of members of a few organizations (e.g., Schall, 1983; Hofstede, et al., 1990; O'Reilly, Chatman and Caldwell, 1991; Chatterjee, et al., 1992). The questionnaires are usually designed to measure important elements of culture that can then be compared across firms to draw conclusions about how they differ in culture and how culture affects organizational performance. These studies are useful in that they provide concrete

empirical measures of differences between firms on several dimensions related to culture. However, there is often little agreement from one investigation to the next concerning the key elements of culture. Moreover, these studies often have small numbers of independent observations (firms) and the usual concerns in survey research like response bias.

Perhaps because of this difficulty of measurement, culture has received considerably less attention from economists. Kreps (1990) argues that culture presents organizations with a solution to problems resulting from multiple equilibria in which there may be uncertainty about the appropriate behavior. (Cultural rules are “focal principles” that point to a socially understood solution, limiting the need for explicit communication.) By having self-enforcing – but not necessarily explicit – rules concerning the appropriate behavior in these types of situations, firms are better able to cope with strategic uncertainty. Hermalin (2000) summarizes the economic approach to culture and presents a formal model in which culture is an efficiency-improving asset in which firms can invest. Arrow (1974) discusses “codes” developed by organizations to help coordinate activity and points out that these codes are path-dependent and may therefore differ greatly between firms, even though each is efficient.

Cremer (1993) builds on Arrow’s concept of codes to define culture in a way similar to what we use in our experiments. He defines culture as “the part of the stock of knowledge that is shared by a substantial portion of the employees of the firm, but not by the general population from which they are drawn” (p. 354).<sup>6</sup> In Cremer’s model, the organization must respond to outside messages in a coordinated manner and this is less

costly to accomplish when the stock of shared knowledge is greater, because the time needed for communication is lower. Our experiments can be seen as ways of creating these focal principles and codes in the laboratory, and then measuring their empirical consequences.

Taken together, previous work on culture points to a couple of key elements. Shared understanding among organizational members is a key element, which usually comes about through shared experience (or a process of socialization and handing-down of traditions). Organizational researchers and economists agree that this shared understanding is useful to the organization because it allows members of a firm to coordinate activity successfully. Our implementation of culture in the laboratory – similar in many ways to Arrow and Cremer's definition – includes these elements and, therefore, allows us to study culture empirically in a controlled and novel way.

The building blocks of our paradigm are minimal analogies to key features of corporate culture. Each pair must learn to jointly react to an external event (the subset of pictures selected by the experimenter) in order to accomplish a specific task in which the employee has to perform the appropriate actions for that particular external event (selecting the correct subset of pictures). Subjects begin by using elaborate descriptions of the pictures that focus on several features. After several rounds their descriptions are honed to a pithy phrase or word that instantly conveys a shared understanding of what is most distinctive about the picture (this is like the development of jargon, distinctive anecdotes, grammars, and so forth (e.g., Schall, 1983; Schein, 1983). While it is possible to create a richer culture experimentally, this minimal culture is all we need to create

---

<sup>6</sup> This definition is similar to the conceptualization underlying the empirical approach of Hofstede, et al., (1990) for identifying culture. Also, Barley (1983) discusses culture as implicit codes for interpreting



possible conflict. If even this minimal aspect of culture results in post-merger difficulties, then we would expect that differences in more elaborate forms of culture would result in potentially even greater conflict.

The picture-matching task is analogous to situations in which workers can perform any of several possible related activities, but the correct one depends on some information held by managers. Managers have to communicate the information they have to employees as quickly and accurately as possible. If a concise natural language to do this does not previously exist, the organization must develop one to perform efficiently.<sup>7</sup>

The pithiest kinds of languages we have in mind are the precise codes which arise in police work, air traffic control, film direction, the military, and so forth.<sup>8</sup> In these enterprises, mental and physical choreography must be done well or else millions of dollars, and many lives, can be lost. Since timing is so important, having a specialized language in which to communicate the flow of desired actions is extremely valuable, and linguistic mistakes are very costly.

The merger phase of our experiments is analogous to the integration of two cultures after a merger, in which a single manager is “fired” to avoid duplication of

---

action in complicated work environments.

<sup>7</sup> Some similar experiments of this sort have been done in game theory recently (e.g., Blume et al, 1998). They study whether players in sender-receiver games with common interests can develop a homemade language which maps privately-observed types into choices of symbols (such as \*, & and #). Language formation works well, although it is undermined substantially when there is conflict of interest between the sender's desire to hide their type and the receiver's interest in figuring it out.

<sup>8</sup> Examples of development and learning of specialized jargon are policemen (figuring out what crime code to use in reporting an incident), lawyers (categorizing cases and applying precedents), film-making (“key grip,” “best boy”), and doctors (medical jargon). Note in each case how long it takes newcomers to learn the language and how important it is for them to do so. Also interesting are businesses in which words are used to describe visual or sensory images that are not ordinarily translated into language-- such as wine tasting, art, and music. (Musicians in recording studios create a rich language to describe in words the kinds of sounds they are striving for-- calling a drum sound “too crunchy”, “too tinny”, “cheesy”, and so forth.

management. The retained manager then must convey instructions to new (acquired) workers, who are used to a different corporate culture, and to the old workers. Our hypothesis is that the merger will fail, in the sense that it will take the manager and workers considerably longer to figure out which pictures to choose.

Merger failure will occur if the culture each pair develops is path-dependent and idiosyncratic. Even if the two firms perform efficiently in the pre-merger period because of their concise languages, if their languages are different then post-merger communication among all parties will be difficult. (Indeed, the more deeply-ingrained firm-specific language is-- and the more efficient the firm-- the harder the integration may be.) After a merger it will take some time to either “train” the new employee to understand the acquiring firm’s language or to develop a new common language that all members of the organization share.

We should note that the development of shared language for naming objects is really a metaphor for broader ideas of culture. Culture is usually defined as a system of values and ideals (what’s good), norms (what’s expected), and conventions of behavior (how things are done). Agreeing on code names for objects is a huge simplification of all that culture is. But as we stress throughout, it is meant to be just a starting point

The difficulty of cultural integration is well known. The more interesting question is whether subjects are aware of the extent of the difficulty associated with merging. If not, they will underestimate the degree of difficulty in resolving cultural conflict. This is measured in a treatment in which subjects make guesses (for money) about how quickly pictures can be matched in the post-merger phase. Treating these

---

They also use analogy if it is likely to be commonly understood, as in "How about that 'In the Air Tonight' sound?" or "More 'Born in the USA', please". )

values as subjects' estimates of the value of the merged firm, we can investigate whether the highest bidder suffers from a degree of "winner's curse,"<sup>9</sup> which has been suggested as an important reason why acquiring firms overpay in mergers (e.g., Roll, 1986). We can also test directly whether the subjects' estimates are more optimistic than the actual performance of the merged firm.

The point of the experiments is not, of course, to capture all the complexity of naturally occurring mergers between large companies; it's precisely the opposite. Like "toy models" used in economic theory to explore logical implications of assumptions and necessary conditions for certain kinds of behavior, the experiments are "toy worlds" which help us think about the essential features of mergers by paring them down. Furthermore, the experimental paradigm we explore is a beginning, not an end. Precisely because the experiments are so simple, one can generate a long list of features of actual business mergers that could, in principle, be added to future experiments to see whether they matter (we will mention a few in the conclusion).

## **Cultural conflict in mergers**

Two examples may help illustrate our ideas about cultural conflict in mergers.

One is the Daimler-Chrysler merger. In the period leading up to the merger, both firms were performing quite well (Chrysler was the most profitable American automaker), and there was widespread expectation that the merger would be successful. People in both organizations expected that their "merger of equals" would allow each unit to benefit from the others capabilities: Chrysler would get the benefit of association

---

<sup>9</sup> In auctions for objects of common unknown value, the winner's curse refers to the fact that bidders fail to

with the Mercedes-Benz name, Daimler-Benz would benefit from Chrysler's highly regarded design staff, and both automakers would benefit from combining purchases and distribution and from access to the other's market. Stockholders in both companies overwhelmingly approved the merger and the stock prices and analyst predictions reflected this optimism. It was expected that both units would operate in the same way that they had prior to the merger and retain most of their own management.

Performance after the merger, however, was entirely different, particularly at the Chrysler division. The stock price has fallen by roughly one half since the immediate post-merger high. Differences in culture between the two organizations are largely responsible for this failure. Operations and management were not successfully integrated as "equals" because of the very different ways in which the Germans and Americans operated: while Daimler-Benz's culture stressed a more formal and structured management style, Chrysler favored a more relaxed, freewheeling style (to which it owed a large part of its pre-merger financial success). As a result of these differences and the German unit's increasing dominance, performance and employee satisfaction at Chrysler took a steep downturn. Following the merger, Chrysler began losing money (and continues to do so – in fact, DaimlerChrysler announced recently that the Chrysler division is not expected to generate profits before 2003). There were large numbers of departures among key Chrysler executives and engineers. DaimlerChrysler is now operated almost exclusively by the German unit, which is increasingly dissatisfied with the performance of the Chrysler division, leading to additional layoffs. Employees at Chrysler are extremely dissatisfied with what they perceive as the source of their

---

adjust for the fact that high bids reflect the most optimistic guesses about the object value (and statistically, will tend to be higher than the object's true value).

division's problems: Daimler's attempts to take over the entire organization and impose their culture on the whole firm (for instance, one joke told at Chrysler is: "How do you pronounce DaimlerChrysler? . . . 'Daimler' the 'Chrysler' is silent.").

Another example is AOL-TimeWarner. A front-page Wall Street Journal article shortly after the merger talked about possible synergies, etc. The only clear discussion of possible cultural conflict is a single paragraph (out of a 60 column-inch article) talking about the "different personalities" of AOL's Steve Case and TimeWarner's Gerald Levin and how their personalities reflect cultural differences. (AOL is more entrepreneurial and flatter; TimeWarner is notoriously hierarchical, status-oriented, and political.)

In a similar website article on MSN's Moneycentral, Jim Jubak included a single paragraph entitled "What could go wrong with the synergy strategy". He wrote:

"Time Warner's steeply hierarchical, highly politicized, and often slow-moving bureaucratic corporate culture could drive crucial divisional managers at America Online nuts and send them screaming out the door. If that happens, goodbye synergy—the people who know how to put Time Warner content on the Internet will be working somewhere else." In these sorts of cursory, obligatory discussions of possible cultural conflict, there is rarely discussion of what steps might be taken if there is dramatic conflict.

While culture may seem like a "small thing" in mergers, compared to product-market and resource synergies, we think the opposite is true because culture is pervasive. It affects how the everyday business of the firm gets done—whether there is shared understanding during meetings and in promotion policy, how priorities are set and whether they are uniformly recognized, whether promises that get made are carried out,

whether the merger partners agree on how time should be spent, who the enemies and friends are, and so forth. And cultural differences may be larger at the top (among senior executives), not smaller.

Many previous studies have touched on aspects of merger failure, though there are none that conclusively document the causal effect of cultural conflict. Most studies simply document success or failure of mergers, without directly addressing differences in culture (e.g., Ravenscraft and Scherer, 1987 & 1989).

Some studies examine the effects on post-merger profitability of product and resource relatedness (Singh and Montgomery, 1987; Shelton, 1988; Harrison, et al., 1991). Most evidence suggests that mergers are more successful when merging firms make related products (usually measured by proximity of SIC code). Other studies also examine similarities or differences in some areas related to corporate culture, without directly addressing culture (Shanley and Correa, 1992).

The studies best able to establish causal effects of conflict simply are not conclusive. For instance, using coders' evaluations of mergers from case studies, Larsson and Finkelstein (1999) show that organizational integration is related to synergy realization (but that it is not related to management style similarity). However, one problem with their study – and other similar studies – is the halo effect, or coding bias that leads synergy realization and organizational integration to be coded as closer together than they really are.

In another study, Chatterjee, et al., (1992) sent questionnaires to top managers of 198 acquired firms, asking them (years after the merger) to rate how different the acquiring and acquired firms were on seven dimensions of cultural orientation (e.g.,

innovation, lateral integration, etc.) and on how strongly the acquiring firm had imposed goals and decisions on the firm. However, of the 198, only 30 firms were used in the analysis, meaning 85% were excluded.<sup>10</sup> Their hypothesis that mergers in which cultural differences were rated to be large would be less successful was supported by their analysis: the cultural difference and imposed-control measures were both highly correlated with stock return residuals from 7- or 16-day windows around the merger announcement date, in an OLS regression of residuals on the two measures and a control variable (relative size). This finding is consistent with the hypothesis that differences in acquiring and acquired firm cultures are bad for merger success (as rated by stock market anticipations). However, one big problem with their study is that the OLS results are overturned by a more conservative chi-square test using median splits on cultural differences and return residuals, which yields no significant difference.<sup>11</sup> In addition, other potential problems with their study include: response bias, survivorship bias (presumably, only acquired firm managers who stuck with the firm returned questionnaires), questions based on respondent memory, a weak measure of actual merger performance (stock market residuals around the merger announcement), and a small sample size.

We do not mean to pick on these studies. Instead, we want to highlight how difficult it is to have an ideal field study and hence, how helpful (at the margin) some exploratory experiments might be.

---

<sup>10</sup> Of the 198 firms, there were only 56 for which two or more managers returned the questionnaire. Including only companies whose merger announcement dates could be clearly identified and whose stock prices were reported further reduced the sample. Catterjee, et al., also excluded firms for which manager responses were not consistent.

<sup>11</sup> Their Table 6 reports that 40% of the 15 high-difference observations are above the median in stock return residuals (correcting for an apparent typo, “below” for “above”), and 60% of the 15 low-difference

Imagine the kind of data we would like to have to test whether cultural differences lead to a surprising degree of merger failure. First, we would have pre- and post-merger measures of culture for both firms, collected independently of any possible merger activity (to avoid endogeneity). We would also have short-run measures of predicted performance (e.g., top manager purchases of acquiring firm shares) and long-run measures of performance (in case the stock market does not forecast accurately). Ideally, it would also be useful to have measures that point to how managers might reduce the risks of failure in the future. As far as we know, such a study does not exist. This is not surprising considering the requirements of the data.

As a result, field data are always partially inconclusive for several reasons. There is no perfect control group (only econometric control of how the acquired firm would most likely have performed without the merger – as in Ravenscraft and Scherer, 1987 & 1989). Standardized measures of culture are difficult to obtain for large samples of firms without some concern about response bias. It is also virtually impossible to have ex-ante data on a large enough sample of firms to randomly pick up those who happen to merge. The alternative is to quickly measure pre-merger culture after a merger is announced (or is rumored), which raises concerns about endogeneity and reporting bias using ex-post recollection (as in Chatterjee, et al, 1992). There is also no way to perfectly isolate the effects of culture in modestly sized samples where there may be year and industry effects, and much idiosyncrasy.

---

observations are above the median in residuals. The Chi-square of 1.2 is not significant at conventional levels, which suggests the OLS result may be entirely driven by influential outliers.



Experiments control for all these concerns. This control comes, of course, at a price – reduced generalizability – which springs from the fact that the experiments lack all the features of the sample of mergers being generalized to.

In sum, field tests have some of the features of the ideal (impossible) study. Experiments have other features. Therefore, the right way to proceed seems to be to do both. Experiments enable us to precisely document whether cultural clashes (of a simple kind) create merger failure (of a simple kind). They also allow us to examine the underlying mechanisms (by manipulation) in the way that field studies rarely do. The main goal of using experiments this way is as a complement to the search for phenomena in the field.

### **Some related experimental research**

There are at least three experimental studies on aspects of organization that are related to this research.

Knez and Camerer (1994, SMJ) studied a 'weak-link game' in which each player chooses numbers 1-7. Each member of the group earns a payoff which is increasing in the minimum number anybody chose, and pays a penalty if their own number is above the minimum. There are many Nash equilibria (any common choice  $X$  is an equilibrium, since  $X$  is a best-response to everyone choosing  $X$ ) but the highest number, 7 is the payoff-dominant equilibrium. Previous research showed that two-person groups could reach the best (efficient) equilibrium and groups of size 6 or more **always** decay to the worst equilibrium (all choosing 1). Groups of size 3 appear to represent a phase transition in a sociodynamic process-- some reach efficiency, acting like pairs; and some tumble

toward inefficiency, acting like large groups. Knez and Camerer asked what happened when two 3-person groups are merged together into a 6-person group. Since larger groups flounder toward inefficiency, the most likely outcome is that the 6-person groups do poorly. Alternatively, the announcement of a merger could conceivably act like a public coordinating device, which tells the group, “We did poorly before in our 3-person groups; let's use this opportunity to do better.” In fact, the 6-person merged groups did poorly, converging to the lower of the two 3-person minima in every case, and to the absolute minimum (1) in 80% of the groups.

Weber, et al.'s, (in press) study of credit and blame attribution began with the observation that there is strong influence of group size in the weak-link games, as described in the preceding paragraph. However, they noted, the distributions of numbers players choose in the first period of the game are very similar across a large range of group sizes (from two to 16). This suggests that the players in very large groups, for example, do not realize how likely it is that in a large group one or more players will choose a low number, and that the large group will eventually end up at an inefficient equilibrium of one. In social psychological terms, the situation (group size) has a stronger influence on behavior than individual dispositions (since even those who began choosing 7 will end up choosing 1 at the end). However, many studies have shown that people tend to over-attribute cause to individual trait differences rather than to situational variables (the “fundamental attribution error”). Weber, et al., therefore conjectured that in small and large groups that reliably converge to efficiency and inefficiency, respectively, players would blame other players-- and especially a prominent “lightning rod” player if one existed-- for any failure or success that is due to the surprising

influence of group size. They created such a lightning rod by having a randomly chosen “leader” make a short speech to the group after they had played one or two rounds (before full equilibration had occurred). In fact, players did over-attribute credit and blame to the leader's speech rather than to the effect of group size. Ratings of the speech quality after several periods of play were downgraded (relative to ratings made immediately after the speech) in large groups, where inefficiency resulted, and slightly upgraded in small groups that achieved efficiency. Furthermore, players put their money where their mouths were-- when they played a second game, they voted (at a cost to themselves) to “fire” the leader in the large-group condition more often than in the small-group condition.

Knez and Camerer (in press) exploited the power of precedent and group size to create 3-person groups which played weak-link games efficiently, by starting with 2-person groups (which reliably reach efficiency) and adding a third person after several rounds, who knew the group's history. Inspired by this observation, Weber (2000) conjectured that large-sized groups which play efficiently could be “grown” from smaller groups by adding small numbers of players, provided the growth path was slow and steady enough that adding new blood did not create too much strategic uncertainty (leading players to choose lower, safer numbers). Weber found in experiments that slow growth could lead to 12-person groups which played much more efficiently than groups who began with 12 people (and invariably converged to playing 1). More interestingly, he found that when some “manager” subjects could endogenously control the growth rate, they usually chose to grow too quickly. After a couple cycles of hyper-growth and subsequent “downsizing”, some players learned to grow slowly but others did not.

These experiments are meant to be illustrative of the sorts of organizational phenomena that could be unpacked experimentally. The Knez-Camerer experiments make the simple point-- an obvious one-- that in production processes which are sensitive to the quality of the worst input (like a chain which is as strong as its weakest-link), combining two small groups into a large one will generally harm performance (absent special steps to limit the harm). The attribution experiments show that it may be common, even in a very simple setting in which a subject randomly chosen to make a speech should hardly be blamed, for players to have an "illusion of leadership", resulting in leader credit and blame for the effects of situations which are easy to manage and hard to manage. The growth experiments show that in weak-link production processes, managers' instincts are to grow firms too swiftly, but controlled growth can yield a degree of efficiency in large groups which is astonishing relative to the baseline results (from groups that are "born large").

A common complaint about these earlier results is that coordination games are often easily resolved by communication-- since players have a common interest-- and players were not allowed to talk (except for the leader's speech). The experiments in this paper circumvent this criticism entirely by making communication the main focus.

## **Methods**

The task in our experiments is based on studies by Clark and Wilkes-Gibbs (1986) and Schober and Clark (1989) to address how shared meaning arises in language. In the experiments, every subject is presented with the same set of 16 pictures depicting office environments (see Figure 1 for examples). While most of the pictures share some

common elements – people, furniture, room characteristics, etc. – each picture is unique. Among the aspects that vary are the number of people and their characteristics (gender, clothing, ethnicity), physical aspects of the room (high ceilings, objects on walls, furniture), and the actions of the people (conversing with others in the picture, talking on the phone, working at a computer).

In every round of the experiment, the experimenter indicated 8 of the 16 pictures in a specific order to one subject, in the role of “manager.” The manager then described the pictures any way he or she liked to the other subject (the “employee”). The employee had to select the correct 8 pictures, in the same order as indicated by the experimenter to the manager. The manager and employee each earned the same amount of money based on how quickly the task was completed. Specifically, the completion time was measured up to the nearest 10 seconds and both subjects earned \$1 minus the completion time divided by 300.

In the first part of each experimental session, two pairs of subjects repeated this task for either 15 or 20 rounds. We expected that managers would initially struggle to convey the information to the employees, requiring long, elaborate descriptions before both subjects could agree on a particular picture. However, with experience they should be able to develop more concise descriptions to refer to each picture and would be able to perform the task more quickly.<sup>12</sup> We expected these descriptions would be idiosyncratic and would only be effective because of the shared history and understanding that would emerge from repeated interaction with the same co-worker.

---

<sup>12</sup> To see precisely how the development of shared understanding takes place in a similar task, see the studies by Clark and Wilkes-Gibbs (1986) and Schober and Clark (1989).



Figure 1. Sample stimuli for experiments.

After allowing the groups many rounds of experience with the task, so they could develop a shorthand language, the two groups were merged. One of the managers was randomly selected to take over both employees and the other manager was paid and left the experiment. The remaining – or acquiring – manager then performed the same task simultaneously with two employees for an additional 10 rounds. In this second part of the experiment, each employee could complete the task individually of the other employee and could therefore receive a different amount in each round for completing the task. The manager received either the average or the sum of the earnings of both employees for the remaining rounds.

There were two slightly different versions of the experiment. In the first, subjects repeated the task for 15 rounds in fixed roles. That is, they were either a manager or employee for the entire 15 rounds. After these 15 rounds, one firm was randomly selected to “acquire” the other and the experiment continued for 10 rounds as described above.

The second version of the experiment was identical except for three differences. First, the task was initially performed for 20 rounds. This was done to allow the groups a little more time to develop and get “locked-in” to their language. A second difference was that the roles in the first part of the experiment were not fixed. Each subject alternated between the roles of manager and employee, resulting in 10 rounds as manager for each participant. This was done to reduce the idiosyncratic effects of a particular manager who might be worse at developing a language. Moreover, by letting both subjects serve in the role of manager we expected they would be more likely to converge on a common language rapidly.

Finally, in the second version of the experiments subjects were also asked to state a belief about the average time it would take the two post-merger employees to complete the time over the remaining 10 rounds. This belief was elicited after the end of the first 20 rounds but before the composition of the post-merger firm was determined. Therefore, subjects had to base their expectation on the performance histories of the two pre-merger firms (subjects had access to the full history of completion times for the other firm as well as for their own). Subjects were told that the one subject who was left out of the post-merger firm would be rewarded for a correct prediction (\$10 for being within 3 seconds of the actual average time and \$4 for being within 20 seconds). This was done to obtain a measure of the expected performance of the post-merger firm (similar to analysts' predictions) that may be biased if the potential culture conflict is not taken into account.

## **Results**

This task creates simple “cultures” by requiring subjects to develop conversational norms to quickly refer to pictures. For instance, one pair of subjects began by referring to a particular picture as: “The one with three people: two men and one woman. The woman is sitting on the left. They’re all looking at two computers that look like they have some PowerPoint graphs or charts. The two men are wearing ties and the woman has short, blond hair. One guy is pointing at one of the charts.” After several rounds, this group’s description of this picture was condensed to simply “PowerPoint.”

Of course, these cultures can be extremely idiosyncratic, because they seize on distinctive shards of language developed in the long initial descriptions, or on shared



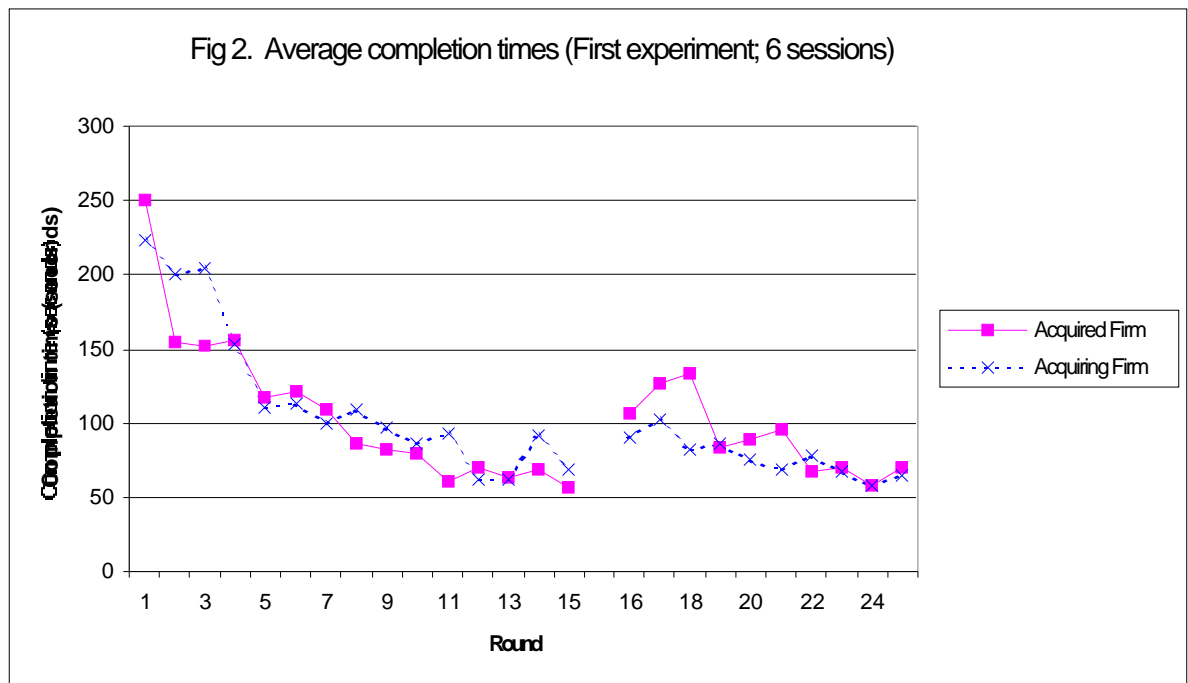
experience within the pair, which is unlikely to be common to outside observers. This process results in sound-bite descriptions that often focus on different aspects of the pictures in different pairs. For instance, the picture called “PowerPoint” by one group was called “Ugly blonde and laptops” by another group. It is unlikely that members of one group would immediately be aware of which picture the other group was referring to by hearing their description.

In another example, one pair of student subjects began referring to a picture as “Uday Rao,” because a person in the picture resembled a professor by that name that taught a class both students were taking. The idiosyncrasy enables a pair to become extremely efficient (cf. Arrow, 1974) but idiosyncrasy also makes it more difficult to combine two pairs and create a common language. In the example, when the new employee was brought in, he had no idea who Uday Rao was, so it took extra time to backtrack and find a different way to describe that picture (and wasted the time of the old employee, who did know who Uday Rao was).

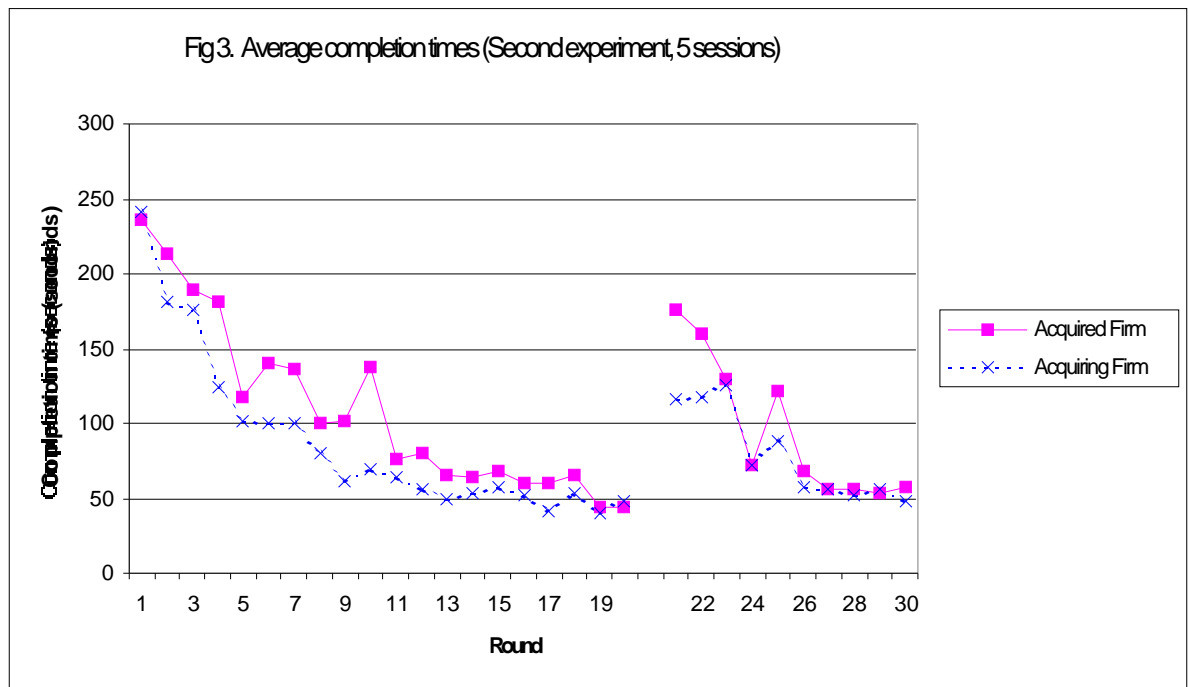
Figures 2 and 3 show the series of completion times across periods for the two types of experiments. The first thing to notice in both figures is that the completion times are initially high: the average completion time in the first round is about 238 seconds in both types of experiments (237 for experiment 1; 239 for experiment 2). The time it takes to complete the task decreases in both figures as groups develop concise ways to refer to the pictures. By the 15<sup>th</sup> round, the average completion time is 63 seconds for both experiments and this goes down to 46 seconds by the 20<sup>th</sup> round in the second type of experiment. Therefore, while the task is initially somewhat difficult because groups lack a common way to refer to the pictures, they become much quicker once they develop

a language. The time series in both periods resemble other types of organizational learning curves (see Argote, 1996).

In both figures, the average completion time increases once the groups are merged. In the first type of experiment, the average completion time increases from 63 seconds in round 15 to 98 seconds in the first post-merger round; while in the second type the average time increases from 46 seconds in round 20 to 146 seconds in the first post-merger round. While in both figures the amount of time it takes to complete the task decreases somewhat with experience, this improvement takes several rounds. While the average completion time in rounds 11-15 (pre-merger) in the first type of experiment is 70 seconds, the merged group is not able to complete the task in this amount of time or less until the 8<sup>th</sup> round (post-merger) and the average completion time in the first five rounds after the merger is 97 seconds.



In the second type of experiment, the merged group is similarly unable to become efficient again quickly. The average completion time for rounds 15-20 in Figure 3 (last 5 pre-merger rounds) is 51 seconds. The merged group is (on average) never able to complete the task in this amount of time or less. The average completion time for the first 5 rounds after the merger is 118 seconds, and the average completion time is 56 seconds in rounds 6-10 after the merger.



As both figures indicate, the merged groups do considerably worse on average than the two separate groups before the merger. This is especially true in the first five periods after the merger, but also continues into later rounds as well. In fact, the outcomes are quite striking: **in all 11 sessions, the average completion times increased from the last five rounds before the merger to the first five rounds after the merger.** Moreover, the average completion time also decreased in all cases between the first five

post-merger rounds and the last five. These outcomes show that language problems create inefficiency in the merged groups even when both firms were performing the task efficiently in the rounds leading up to the merger.

In the second type of experiment, we also collected subjects' estimates of the post-merger completion times. Subjects could receive a bonus if they accurately predicted the average completion time in their session. The average completion time over the last ten (post-merger) rounds in these sessions was 87 seconds, while the average estimate was 72 seconds, indicating that subjects were on average overly optimistic about the amount of time it would take the merged group to complete the task.<sup>13</sup> This difference is significant at the 0.05 level ( $t_{23} = 1.726$ , one-tailed).

Another result from these experiments has to do with conflict between members of the acquiring "firm" and the new employee. One possible source of merger failure is exactly this type of conflict (see Buono, Bodwitch and Lewis, 1985). The first piece of evidence that this occurred in our experiments is anecdotal. In some sessions, there was tension and hostility between the manager of the merged firm and the new employee. This usually resulted from the manager having to provide longer and more detailed explanations to the new employee than to the "experienced" employee. In some cases, these longer explanations still didn't work because the two subjects were focusing on different aspects of pictures. For instance, in one case, the manager was trying to describe a picture which subjects in the acquiring firm had come to refer to as "coffee mugs" because of their presence on a table at the front of the picture. The other group

---

<sup>13</sup> Research by psychologists has shown that people generally under-estimate the time it will take to complete tasks (e.g., Byram, 1997; Buehler, Griffin and Ross, 1994). However, this bias is found when estimates concern unfamiliar tasks and is eliminated when subjects use past experience in informing their

had referred to this picture as “cupboard in back”. The manager and new employee spent close to a minute trying to jointly identify this picture. During this time, the manager would repeat some aspects of the picture (such as the number of people) and then refer to the coffee mugs on the table. The delay was due mainly to the employee taking a while to identify a likely picture (there were other pictures with coffee mugs), but even after she had the correct picture, she tried to verify this by asking if the picture had a cupboard in the back. Even though it did, the manager (annoyed at the delay by now) replied, “I don’t know, just look for the coffee mugs!” leading to a longer delay before the employee could be sure they were talking about the same picture. In another case, an acquired employee angrily interrupted a manager: “Stop telling me what they’re wearing and just tell me how many people are in the picture!” This type of mild hostility resulting from different perspectives occurred in a few cases in the experiment.

A more precise measure of the conflict can be seen in responses to questionnaires that were administered both immediately after the first (pre-merger) part of the experiment and at the end of the ten post-merger rounds. The questionnaire included items asking people to evaluate whether the other participants are “better or worse at this task than the average CMU [or Caltech] student would be?” Of particular interest, subjects who were acquired in the merger gave the manager they worked with an average rating of 6.6 on a 9-point scale. These same employees, however, gave the new manager only a 5 ( $p < 0.10$ ,  $t_{12} = 1.577$ , one-tailed). Similarly, both employees of the acquired firm rated the new employee as worse than the employee they were more familiar with (manager: 6.3 vs. 6.6; employee: 6 vs. 7; neither difference is significant). These results

---

estimates. In our case, subjects have both recently participated in 20 rounds of the task and are given the full list of previous completion times when making their estimates.

are consistent with the attributions literature in which people over-assign responsibility for outcomes to others' personal traits, relative to situational variables (Weber, et al., in press; Ross and Nisbett, 1991). While the simple questions used here do not constitute a behavioral measure, Weber, et al., (in press) show that a similar result holds when subjects take actions that determine monetary incentives.

## **Conclusion**

The idea guiding this research is that failures to coordinate activity, based on cultural conflict, contribute to the widespread failure of corporate mergers. Furthermore, we suggest that the likelihood of coordination failures is underestimated, which explains why firms enter into so many mergers that are culturally doomed in the first place. Our experiments support both hypotheses: differences in culture between our laboratory firms lead to consistent decreased performance for both employees after the merger, and subjects under-predict the extent of this decrease. In addition, we also find evidence of conflict arising from the differences in culture, pointing to a possible source for the high turnover rate following real mergers.

Future work might deal with omissions from these experiments. For instance, a likely omission is conflict of interest or perspective. For example, suppose one pair represents "human resources" and is trained to describe pictures in terms of their human features; another represents "technology" and talks in terms of machinery. It is possible that this specialization in function and skill leads to a high degree of specialization in language-- and a blindness towards language of others, which makes post-merger communication even more difficult. This is consistent with work by Dearborn and Simon

(1958) and Dougherty (1992) indicating that specialists may have a difficult time taking the perspective of people in other areas.

Other ideas we would like to incorporate (some of these would involve adjusting the paradigm to allow for larger “firms”):

Decayed organizational memory and the evolution of culture: In our experiments, the retained manager remembers how difficult it was to create language initially (in the pre-merger phase). This may give him or her extra patience or skill in creating a new hybrid language (or explaining the old language) when the merger occurs. A twist on our design is to bring in additional subjects during only the last few trials of the pre-merger phase, when the language is brief and working well. They may be more prone to overestimate how easy the post-merger transition will be, since they have not experienced the worst part of the early pre-merger learning. Another possibility is that bringing in new people to replace old employees may maintain the organization’s culture, but make it especially susceptible to ignoring how difficult it is to develop a language.

We are also interested in exploring how organizational culture may evolve or change as new people are brought in to the organization to replace existing members. In particular, is it possible for the organizational culture to remain intact even once the original employees have all been replaced? Another issue is how the development of culture is affected by the patterns of interaction between members of the organization.

Turnover and blame: Questionnaires we administered after the experiments suggest a typical pattern of blame consistent with research on attribution errors: managers blamed the new employee, and new employees blamed the managers, for slow post-merger communication. It would therefore be interesting to allow managers to

terminate employees, and allow employees to quit. One hypothesis is that they would terminate and quit too hastily, not anticipating how much easier communication becomes with practice. Hasty termination could backfire for managers if the new employees who replace the fired ones take even longer to learn the new language. On the other hand, it is conceivable that the replacements would do better learning the manager's language since they start tabula rasa.

Solutions: If we can produce reliable problems with post-merger integration, then we can also test possible solutions. For example, in the experiments a couple of subjects asked, "Can we just have a few minutes to talk about this?" before beginning the merger phase of the experiment. A consultant at McKinsey told us that his prescription for improving post-merger integration is to create a new task on which employees from both the acquiring and acquired firm work together. By using a new task, the employees are inhibited from using the full extent of the language that is familiar from old tasks, and are able to compromise on a new shared language. The effectiveness of prescriptions like these could be easily tested in experiments.

While the above all point to additional work that will hopefully produce additional insights into how mergers and culture operate in the real world, our main point in this paper was to use a simple experimental procedure to make an important point. By recreating a simple form of culture in the laboratory encompassing many key aspects of real-world organizational culture, we provided a clear situation in which merger failure is driven by differences in this culture.





## References

- Argote, Linda. 1996. "Organizational learning curves: persistence, transfer and turnover." *International Journal of Technology Management* 11: 759-769.
- Arrow, K. A. 1974. *The Limits of Organization*. New York: W. W. Norton & Co.
- Barley, S. R. 1983. "Semiotics and the study of occupational and organizational cultures." *Administrative Science Quarterly*, 28: 393-413.
- Blume, A., D. V. DeJong, Y. G. Kim and G. B. Sprinkle. 1998. "Experimental evidence on the evolution of meaning of messages in sender-receiver games." *American Economic Review*, 88(5): 1323-1340.
- Buono, A. F., J. L. Bowditch and J. W. Lewis. 1985. "When cultures collide: The anatomy of a merger." *Human Relations*, 38: 477-500.
- Chatterjee, S., M. H. Lubatkin, D. M. Schweiger and Y. Weber. 1992. "Cultural differences and shareholder value in related mergers: Linking equity and human capital." *Strategic Management Journal*, 13: 319-334.
- Clark, H. H. and D. Wilkes-Gibbs. 1986. "Referring as a collaborative process." *Cognition*, 22(1): 1-39.
- Cremer, Jacques. 1993. "Corporate culture and shared knowledge." *Industrial and Corporate Change*, 2(3): 351-386.
- Dearborn, D.C. and H. A. Simon. 1958. "Selective perception: A note on the departmental identification of executives." *Sociometry*, 21: 140-144.
- Dougherty, D. 1992. "Interpretive barriers to successful product innovation in large firms." *Organization Science*, 3(2): 179-202.
- Harrison, J. S., M. A. Hitt, R. E. Hoskisson, R. D. Ireland. 1991. "Synergies and post-acquisition performance: Differences versus similarities in resource allocations." *Journal of Management*, 17(1): 173-190.
- Hermailin, Benjamin E. 2000. "Economics and corporate culture." Unpublished manuscript.
- Hofstede, G., B. Neuijen, D. D. Ohayv and G. Sanders. 1990. "Measuring organizational cultures: A qualitative and quantitative study across twenty cases." *Administrative Science Quarterly*, 35: 286-316.
- Knez, M. and C. F. Camerer. 1994. "Creating expectational assets in the laboratory: Coordination in weakest-link games." *Strategic Management Journal*, 15: 101-119.

Knez, M. and C. F. Camerer. In press. "Increasing cooperation in social dilemmas by establishing a precedent of efficiency in coordination games." *Organizational Behavior and Human Decision Processes*.

Kreps, David M. 1990. "Corporate culture and economic theory." In *Perspectives on Positive Political Economy*, eds. J. E. Alt and K. A. Shepsle. Cambridge, England: Cambridge University Press.

Larsson, R. and S. Finkelstein. 1999. "Integrating strategic, organizational, and human resource perspectives on mergers and acquisitions: A case survey of synergy realization." *Organization Science*, 10(1): 1-26.

Marcoulides, G. A. and R. H. Heck. 1993. "Organizational culture and performance: Proposing and testing a model" *Organization Science*, 4(2): 209-225.

McGuckin, R., S. V. Nguyen and A. P. Reznick. 1995. "The impact of ownership change on employment, wages, and labor productivity in U.S. manufacturing 1977-87." Center for Economic Studies Working Paper, CES 95-8.

O'Reilly, C. A., III., J. Chatman and D. F. Caldwell. 1991. "People and organizational culture : A profile comparison approach to assessing person-organization fit." *Academy of Management Journal*, 34(3): 487-516.

Ravenscraft, D. J. and F. M. Scherer. 1987. "Life after takeover." *Journal of Industrial Economics*, 36(2): 147-156.

Ravenscraft, David J. and F. M. Scherer. 1989. "The Profitability of Mergers." *International Journal of Industrial Organization*, 7(1): 101-116.

Roll, Richard. 1986. "The hubris hypothesis of corporate takeovers." *Journal of Business*, 59(2): 197-216.

Ross, L., R. E. Nisbett. 1991. *The Person and the Situation*. New York: McGraw-Hill.

Rousseau, D. 1990. "Assessing organizational culture: The case for multiple methods." In *Climate and Culture*, ed. B. Schneider. San Francisco: Jossey-Bass.

Schall, Maryan S. 1983. "A communication-rules approach to organizational culture." *Administrative Science Quarterly*, 28: 557-581

Schein, Edgar H. 1983. "The role of the founder in creating organizational culture." *Organizational Dynamics*, 12: 13-28.

Schein, Edgar H. 1990. "Organizational culture." *American Psychologist*, 45(2): 109-119.

Schein, Edgar H. 1996. "Culture: The missing concept in organization studies." *Administrative Science Quarterly*, 41(2): 229-240.

Schober, Michael F. and Herbert H. Clark. 1989. "Understanding by addressees and overhearers." *Cognitive Psychology*, 21(2): 211-232.

Shanley, M. T. and M. E. Correa. 1992. "The agreement between top management teams and expectations for post acquisition performance. *Strategic Management Journal*, 13: 245-266."

Shelton, L. M. 1988. "Strategic business fits and corporate acquisition: Empirical evidence." *Strategic Management Journal*, 9: 279-287.

Singh, H. and C. A. Montgomery. 1987. "Corporate acquisition strategies and economic performance." *Strategic Management Journal*, 8: 377-386.

Walsh, J. P. 1988. "Top management turnover following mergers and acquisitions." *Strategic Management Journal*, 9: 173-183.

Walsh, J. P. and J. W. Ellwood. 1991. "Mergers, acquisitions, and the pruning of managerial deadwood." *Strategic Management Journal*, 12: 201-217.

Wilkins, A. L. and W. G. Ouchi. 1983. "Efficient cultures: Exploring the relationship between culture and organizational performance." *Administrative Science Quarterly*, 28: 468-481.

Weber, R. A. 2000. "Organizational growth and coordination problems: An experimental study." Department of Social and Decision Sciences Working Paper, Carnegie Mellon University.

Weber, R. A., C. F. Camerer, Y. Rottenstreich and M. Knez. In press. "The illusion of leadership: Misattribution of cause in coordination games." *Organization Science*.